Multiple regression analysis of big weather-based outage data

Author: Boyang Ren
Department of Mathematical Sciences
Indiana University-Purdue University Indianapolis

Big data analytics has been recently used in various fields in order help derive data-driven knowledge. In this study, we use statistical methods and data analytics techniques to investigate the potential relationship between weather conditions and power outages. Outages are not only a major source of inconvenience for the community but they also have a major impact on the local economy. In this study, we show that it is not possible to predict outages simply from past history and we investigate whether certain weather factors can help identify the potential of an outage pattern. Our investigation focuses on three regions in northern Indiana, namely, Gary, Goshen, and Valparaiso. These regions were selected as a representative set of different types of communities, namely industrial, urban and rural areas.

There are several technical causes for outages, our study excludes those and focuses on outages that have the potential to be impacted by weather conditions including: a) Tree falling due to storm, b) Tree falling naturally, c) Tree growth naturally, d) Weather and e) Car hitting pole. Outage causes a) and d) are clearly related to the weather conditions. The key is to determine which weather conditions have a direct impact on these causes. We also investigate the potential relation between causes b), c) and e) and weather conditions in the area. The weather conditions that are considered in the study include: precipitation, snow depth, visibility, wind gust, and wind speed. Our study also shows that the number of outages varies from one season to the next and it is mainly influenced by precipitation. Furthermore, the outages due to Car hitting pole is decreasing from year to year probably due to improved safety conditions.